

C:Superfamily: hexokinase; hexokinase homology

Query Match 34.4%; Score 52; DB 2; Length 502;
Best Local Similarity 41.7%; Pred. No. 9.4;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 8 VALCLSPPLAARRRRSSGFFVA 31
DB 14 VAVCAAAALIVRRRMRKAGKAWRY 37

RESULT 3

hypothetical protein PA1232 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83492
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardi, K.; Lim,
L.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MIMD:20457337
A:Accession: G83492
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-679 <STO>
A:Cross-references: GB:AE004552; GB:AE004091; NID:g9947150; PIDN:AA604621.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics: PA1232

Query Match 33.8%; Score 51; DB 2; Length 679;
Best Local Similarity 43.3%; Pred. No. 18;
Matches 13; Conservative 3; Mismatches 8; Indels 6; Gaps 1;

OY 1 MALRLNDVALCLSPPLAARRRRSSGFFVA 30
DB 620 IAVRHRLD-----PPLAARRRRRAGARLAA 643

RESULT 4

hypothetical protein SPBC27B12.06 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40030
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z21900
A:Accession: T40030
A:Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-918 <WOO>
A:Cross-references: EMBL:AL021766; PIDN:CAA16901.1; GSPDB:GN00067; SPDB:SPBC27B12.06
A:Experimental source: strain 972h-; cosmid c27B12
C:Genetics:
A:Gene: SPDB:SPBC27B12.06
A:Map position: 2

Query Match 33.8%; Score 51; DB 2; Length 918;
Best Local Similarity 45.5%; Pred. No. 24;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 5 LNDVALCLSPPLAARRRRSSG 26
DB 53 LNDGSLCENPADVREKKNSSG 74

RESULT 5

A71568
hypothetical protein CT016 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999

C:Accession: A71568
R:Stephens, R.S.; Kaiman, S.; Lammell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A:Reference number: A71570; MIMD:99000809
A:Accession: A71568

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <ARN>
A:Cross-references: GB:AE001276; GB:AE001273; NID:g3328399; PIDN:AAC67606.1; PID:g332
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics: CT016

Query Match 33.1%; Score 50; DB 2; Length 242;
Best Local Similarity 42.1%; Pred. No. 9;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

OY 1 MALRLNDVALCLSPPLAAR 19
DB 1 MKVXINDQFICISPIYSAR 19

RESULT 6

Immediate-early protein pip92 - mouse

A:Alternate names: cycloheximide-induced protein chx1; proline-rich induced protein 9
C:Species: Mus musculus (house mouse)
C:Date: 26-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 05-Nov-1999
C:Accession: A54722; A58370; A35050
R:Lalnick, B.V.; Lau, L.F.
J. Biol. Chem. 269, 23163-23170, 1994
A:Title: Transcriptional activation of the immediate early gene pip92 by serum growth
A:Reference number: A54722; MIMD:94365016
A:Accession: A54722
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-221 <LAT>
A:Cross-references: GB:L26490; NID:g556318; PIDN:AA00120.1; PID:g556319
R:Charles, C.H.; Simskey, J.S.; O'Brien, T.P.; Lau, L.F.
Mol. Cell. Biol. 10, 6769-6774, 1990
A:Title: Pip92: a short-lived, growth factor-inducible protein in BALB/c 3T3 and PC12
A:Reference number: A36370; MIMD:91061787
A:Accession: A36370
A:Molecule type: mRNA
A:Residues: 1-221 <CHA>
A:Cross-references: GB:M59821; GB:M3756; NID:g200354; PIDN:AAA39931.1; PID:g200355
R:Colclough, C.; Kuhn, L.; Lefkowitz, I.
Proc. Natl. Acad. Sci. U.S.A. 87, 1753-1757, 1990
A:Title: Regulation of mRNA abundance in activated T lymphocytes: identification of n
A:Reference number: A35050; MIMD:90175375
A:Accession: A35050
A:Molecule type: mRNA
A:Residues: 1-26, 'D', 28-58, 'S', 60-221 <COL>
A:Cross-references: GB:M31042; NID:g200759; PIDN:AAA40059.1; PID:g200760
C:Genetics: pip92
A:Gene: pip92
A:Introns: #status absent
C:Keywords: immediate-early protein

Query Match 32.8%; Score 49.5; DB 2; Length 221;
Best Local Similarity 50.0%; Pred. No. 9.8;
Matches 12; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

OY 3 LRLNDV-ALCLSPPLAARRRRSS 25
DB 101 LRVGETPALCDPPPARVSRKRSS 124

RESULT 7

F75393
hypothetical protein - Deinococcus radiodurans (strain R1)

```

A:Cross-references: EMBL:AL079345; PIDN:CA845354.1; GSPDB:GN00070; SCOEDB:SC68.18C
A:Experimental source: strain A312)
C:Genetics:
A:Gene: SCOEDB:SC68.18C

Query Match          32.5%; Score 49; DB 2; Length 258;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 8; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 13 SPPLAARRRRSSC 26
      :|||||::|||:|
Db 229 APVAAKKQQRSTG 242

RESULT 10
AG2330
C:Species: Anabaena sp.
A:Note: Anabaena sp. (Strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
A:Accession: AG2330
R:Nakazaki, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamato, S.; Watanabe, A.; Iriig
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MIMD:21595285; PMID:11759840
A:Accession: AG2330
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <RUR>
A:Cross-references: GB:BA000019; PIDN:BA875897.1; PID:91713333; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: rpl15
C:Superfamily: Escherichia coli ribosomal protein L15

Query Match          31.8%; Score 48; DB 2; Length 148;
Best Local Similarity 39.3%; Pred. No. 11;
Matches 11; Conservative 6; Mismatches 7; Indels 4; Gaps 1;

OY 3 LRLNDVALCLSPPLAARRRRSSGRFVA 30
      :||||| |::||| ||::
Db 1 MRLNDV---KPKQGSKKRRRRVGRGIS 24

RESULT 11
AC2994
C:Species: Agrobacterium tumefaciens
A:Note: conserved hypothetical protein Atu3557 [imported] - Agrobacterium tumefaciens (strain
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
A:Accession: AC2994
R:Wood, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Mo
eraga, G.; Gillet, W.; Grant, C.; Guenhiner, D.; Kutayavin, T.; Levy, R.; Li, M.; McC
e Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Ka
ster, E.W.
A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AC2994
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <RUR>
A:Cross-references: GB:AE008669; PIDN:AAL44369.1; PID:917741964; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3557
A:Map position: linear chromosome

Query Match          31.5%; Score 47.5; DB 2; Length 507;
Best Local Similarity 44.0%; Pred. No. 43;
Matches 11; Conservative 4; Mismatches 7; Indels 3; Gaps 1;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 11:07:05 : Search time 10 Seconds

(without alignments)
120.031 Million cell updates/sec

Title: US-09-358-321c-15

Perfect score: 151

Sequence: 1 MALRLNDVALCLSPPLARRRRSSGRFVAV 31

Scoring table: BLOSUM62

Gapop 10.0, Capext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	52	34.4	502	1 HKX2_ARATH	P93834 arabidopsis
2	49.5	32.8	221	1 IER2_MOUSE	P17950 mus musculus
3	49	32.5	162	1 DSB8_NEISB	Q919C6 neisseria m
4	48	31.8	992	1 AXN1_MOUSE	Q35625 mus musculus
5	46.5	30.8	739	1 UL47_HSVBC	P36338 bovine hepr
6	46.5	30.8	742	1 UL47_HSVBC	P36338 bovine hepr
7	46	30.5	162	1 DSB8_NEISB	Q919C6 neisseria m
8	46	30.5	231	1 PSBO_ONOVI	Q25291 onobrychis
9	45	29.8	147	1 RL15_SYNY3	P73303 synchocyst
10	45	29.8	228	1 VCOM_ADEMI	Q10442 mouse adeno
11	45	29.8	268	1 YGDL_ECOLI	Q46927 escherichia
12	45	29.8	533	1 YD3D_SCHPO	Q10277 schizosach
13	45	29.8	739	1 HYPF_RHOCA	Q02887 rhodobacter
14	45	29.8	3011	1 POLG_HCV1	P26654 h genome po
15	44.5	29.5	246	1 YNCA_YEAST	P53670 saccharomyc
16	44.5	29.5	632	1 NODQ_RHTR	P52978 r nodq bifu
17	44	29.1	377	1 DAPR_COILA	Q39535 coix lachy
18	44	29.1	385	1 FOX1_YEAST	P34161 saccharomyc
19	44	29.1	410	1 YLJ8_CAEEL	P34571 caenorhabd
20	44	29.1	420	1 IE68_HSV1	Q04485 herpes simp
21	44	29.1	468	1 SELA_PSEAE	Q9H01 pseudomonas
22	44	29.1	496	1 HKX1_ARATH	Q4325 arabidopsis
23	44	29.1	891	1 MAX1_SCHCO	P37935 schizophyll
24	44	29.1	3010	1 POLG_HCV1	P26654 h genome po
25	43.5	28.8	137	1 VIT_STRPU	P19615 strongyloce
26	43.5	28.8	641	1 NODQ_RHTR	P52978 r nodq bifu
27	43.5	28.8	1826	1 SUIS_HUMAN	P14410 homo sapien
28	43	28.5	115	1 HUS3_MYCTU	P08539 saccharomyc
29	43	28.5	471	1 GBAL_YEAST	P52638 saccharomyc
30	43	28.5	503	1 VP57_BDV	P29358 h genome po
31	43	28.5	1189	1 ALA8_ARATH	Q91K80 arabidopsis
32	43	28.5	3011	1 POLG_HCV1	P26654 h genome po
33	42.5	28.1	189	1 YC68_ARCFU	Q29000 archaeglob

34	42.5	28.1	405	1 PE21_MOUSE	P35375 mus musculus
35	42.5	28.1	405	1 PE21_MOUSE	P35375 mus musculus
36	42.5	28.1	633	1 NODQ_RHTR	P52978 r nodq bifu
37	42	27.8	57	1 P10304_BPT3	P10304 bacterioph
38	42	27.8	73	1 RS27_AERPE	Q9YF01 aeropyrum p
39	42	27.8	116	1 UIC9_HOMVA	P16838 human cytom
40	42	27.8	208	1 COAT_PURV1	P17522 potato leaf
41	42	27.8	208	1 COAT_PURV1	P17522 potato leaf
42	42	27.8	208	1 COAT_PURV1	P17522 potato leaf
43	42	27.8	209	1 YNU1_SHIFL	P11624 potato leaf
44	42	27.8	238	1 YPE2_RHTR	P29770 shigella fl
45	42	27.8	313	1 MPO1_ALCEU	P23139 rhodospiril
					P17295 alcaligenes

ALIGNMENTS

RESULT 1	ID	HKX2_ARATH	STANDARD	PRT	502 AA.
AC	P93834				
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Hexokinase 2 (EC 2.7.1.1).				
CN	HKX2 OR AT2G19860 OR F6F22.11.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid 11; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV, LANDSBERG, ERECTA;				
RX	MEDLINE=9716664; PubMed=9014361;				
RA	Yang J.C., Leon P., Zhou L., Sheen J.;				
RT	"Hexokinase as a sugar sensor in higher plants.";				
RL	Plant Cell 9:5-19(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV, COLUMBIA;				
RX	MEDLINE=20083487; PubMed=10617197;				
RA	Lin X., Kaul S., Rounsley S.D., Shao T.P., Benito M.-I., Town C.D.,				
RA	Fuji I.C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,				
RA	Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,				
RA	Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,				
RA	Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Cressy T.H.,				
RA	Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,				
RA	Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,				
RA	Venter J.C.;				
RT	"Sequence and analysis of chromosome 2 of the plant Arabidopsis				
RT	thaliana.";				
RL	Nature 402:761-768(1999).				
CC	- CATALYTIC ACTIVITY: ATP + D-hexose = ADP + D-hexose 6-phosphate.				
CC	- SUBCELLULAR LOCATION: CHLOROPLAST OUTER ENVELOPE; CYTOPLASMIC				
CC	SIDE (BY SIMILARITY)				
CC	- SIMILARITY: BELONGS TO THE HEXOKINASE FAMILY.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; U28215; AAB49911.1; -				
DR	EMBL; AC005169; AAC62130.1; -				
DR	HSSP; Q26609; 1BDG.				
DR	Mendel; 7096; Arab1512.7096.				
DR	InterPro; IPR001312; Hexokinase.				
DR	Pfam; PF00349; Hexokinase_1.				
DR	PRINTS; PR00475; HEXOKINASE.				

DR ProDom: PD001109; Hexokinase; 1.
 DR ProSITE: PS00378; HEXOKINSES; 1.
 KW Transferase; Kinase; Glycolysis; ATP-binding; Transmembrane;
 RN Chloroplast.
 FT TRANSMEM 4 24 POTENTIAL.
 FT NP_BIND 101 106 ATP (POTENTIAL).
 FT DOMAIN 171 197 GLUCOSE-BINDING (POTENTIAL).
 SQ SEQUENCE 502 AA; 54489 MW; FFAF0B650D7531C9 CRC64;
 Query Match 34.4%; Score 52; DB 1; Length 502;
 Best Local Similarity 41.7%; Pred. No. 3.1;
 Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
 OY 8 VALCSPPLAARRRRSSGRFVAV 31
 DB 14 VAVCAAAALIVRRMKRSAGKARV 37

RESULT 2
 IER2_MOUSE STANDARD: PRT; 221 AA.
 AC P17950; 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE T-lymphocyte activated protein (Cycloheximide-induced) (CHX1)
 DE (Immediate early response 2 protein).
 GN IER2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90175375; PubMed=2308934;
 RA Coleclough C., Kuhn L., Leikovits I.;
 RT "Regulation of mRNA abundance in activated T lymphocytes:
 RT Identification of mRNA species affected by the inhibition of protein
 RT synthesis.";
 RT Proc. Natl. Acad. Sci. U.S.A. 87:1753-1757(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M31042; AAA40059.1; -;
 DR PIR: A35050; A35050.
 DR MGD: MGI:104815; Ier2.
 FT SIMILAR 57 122 TO MOUSE JUN-D PROTEIN (AA 212-270).
 SQ SEQUENCE 221 AA; 24501 MW; 02CF9C18416A526F CRC64;
 Query Match 32.8%; Score 49.5; DB 1; Length 221;
 Best Local Similarity 50.0%; Pred. No. 3.1;
 Matches 12; Conservative 3; Mismatches 8; Indels 1; Gaps 1;
 OY 3 LRINDV-ALCISPLAARRRRSS 25
 DB 101 LRQETPALCDPPPARVSRKRRSS 124

OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=2017575; PubMed=10710307;
 RA Retlein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwin M.L., Deboy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Maon T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Mastigani V., Pizzo M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58".
 RL Science 287:1809-1815(2000).
 CC -!- FUNCTION: REQUIRED FOR DISULFIDE BOND FORMATION IN SOME
 CC PERIPLASMIC PROTEINS. ACTS BY OXIDIZING THE DSB PROTEIN (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (by similarity).
 CC -!- SIMILARITY: BELONGS TO THE DSB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE005215; AAF41998.1; -;
 DR TIGR: NM81649; -;
 DR InterPro: IPR003752; DsbB.
 DR Pfam: PF02600; DsbB; 1.
 KW Oxidoreductase; Redox-active center; Electron transport; Chaperone;
 KW Transmembrane; Inner membrane; Complete proteome.
 FT TRANSMEM 1 8 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 9 29 POTENTIAL.
 FT DOMAIN 30 35 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 36 56 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 57 64 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 65 85 POTENTIAL.
 FT DOMAIN 86 133 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 134 154 POTENTIAL.
 FT TRANSMEM 155 162 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 35 38 REDOX-ACTIVE (BY SIMILARITY).
 FT DISULFID 101 128 REDOX-ACTIVE (BY SIMILARITY).
 SQ SEQUENCE 162 AA; 17701 MW; CC02718EC7166B77 CRC64;
 Query Match 32.5%; Score 49; DB 1; Length 162;
 Best Local Similarity 52.2%; Pred. No. 2.6;
 Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 OY 9 ALCTSPPLAARRRRSSGRFVAV 31
 DB 49 ALCTAIVLMCRPRRRRAGGLFGAV 71

RESULT 4
 AXN1_MOUSE STANDARD: PRT; 992 AA.
 ID AXN1_MOUSE
 AC O35625; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Axin 1 (Axin inhibition protein 1) (Fused protein) (Fragment).
 GN AXIN1 OR AXIN OR FU.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=9373830; PubMed=9230313;
RA Zeng L., Fagotto F., Zhang T., Hsu W., Vasicek T.J., Perry W.L. III,
RA Lee J.J., Tilgman S.M., Gumbiner B.M., Constantini F.;
RT "The mouse fused locus encodes Axin, an inhibitor of the Wnt signaling
RT pathway that regulates embryonic axis formation.";
RL Cell 90:181-192(1997).
CC -1- FUNCTION: INHIBITOR OF THE WNT SIGNALING PATHWAY. DOWN-REGULATES
CC BETA-CATENIN. PROBABLY FACILITATE THE PHOSPHORYLATION OF BETA-
CC CATEININ AND APC BY GSK-3B (BY SIMILARITY).
CC -1- SUBUNIT: INTERACTS WITH GSK-3B AND BETA-CATENIN. THE INTERACTION
CC BETWEEN AXIN AND BETA-CATENIN OCCURS VIA THE ARMADILLO REPEATS
CC CONTAINED IN BETA-CATENIN. TERNARY COMPLEX. ALSO BINDS TO
CC PLAKOGLOBIN (GAMA-CATENIN), APC, DVL AND P2A (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC STEM CELLS.
CC -1- DEVELOPMENTAL STAGE: WIDELY EXPRESSED AT E10.5 TO E16.5 DAY.
CC -1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B AND DEPHOSPHORYLATED BY
CC P2A.
CC -1- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 DIX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF009011; AAC53285.1; -.
DR HSBP; P49799; INGR.
DR MGD; MG1:1096327; Axin.
DR InterPro: IPR001158; DIX.
DR InterPro: IPR000342; RGS.
DR Pfam: PF00778; DIX; 1.
DR Pfam: PF00615; RGS; 2.
DR PRINTS: PR01301; RGS/PROTEIN.
DR PRODOM: PD001580; RGS; 1.
DR PRODOM: PD003639; DIX; 1.
DR SMART: SM00021; DAX; 1.
DR SMART: SM00315; RGS; 1.
DR PROSITE: PS50132; RGS; 1.
DR Developmental protein: Phosphorylation; Alternative splicing.
KV NON_TER 1
FT DOMAIN 10 18 POLY-ALA.
FT DOMAIN 217 340 RGS.
FT DOMAIN 477 561 GSK-3B BINDING SITE (BY SIMILARITY).
FT DOMAIN 562 630 BETA-CATENIN BINDING SITE (BY
FT SIMILARITY).
FT DOMAIN 910 992 DIX.
FT VARSPLIC 860 895 MISSING (IN ISOFORM 2).
SQ SEQUENCE 992 AA; 109917 MW; 70EBB53D87BD26F CRC64;
Query Match 31.88; Score 48; DB 1; Length 992;
Best Local Similarity 50.04; Pred. No. 25;
Matches 11; Conservative 2; Mismatches 5; Indels 4; Gaps 1;
OY 9 ALCLSP-----LAARRRRSSG 26
DB 43 AACLAPPGHSPRRRRRROG 64
RESULT 5
UL47_HSVBC STANDARD; PRT; 739 AA.
AC P36338;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 80.7 kDa alpha trans-inducing protein (VP8 tegument protein) (VP7
DE protein) (107 kDa protein).
GN VP8 OR UL47.
OS Bovine herpesvirus type 1 (strain Cooper).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COOPER / 34;
RX MEDLINE=93057366; PubMed=1331296;
RA Leboissetre S., Trudel M., Simard C.;
RT "Characterization and transcript mapping of a bovine herpesvirus type
RT 1 gene encoding a polypeptide homologous to the herpes simplex virus
RT type 1 major tegument proteins VP13/14.";
RL J. Gen. Virol. 73:2941-2947(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97164286; PubMed=9010999;
RA Schwytzer M., Steyer D., Vogt B., Lowery D.E., Simard C.,
RA Leboissetre S., Mista V., Vitek C., Paces V.;
RT "Gene contents in a 31-kb segment at the left genome end of bovine
RT herpesvirus-1.";
RL Vet. Microbiol. 53:67-77(1996).
CC -1- FUNCTION: MODULATOR OF ALPHA-TIF (VIM65 PHOSPHOPROTEIN) TRANS-
CC ACTIVATION. UL47 MAY HAVE KINASE ACTIVITY.
CC -1- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
CC HHV-1 13, AND VZV 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z54206; CA90922.1; -.
DR EMBL: M84468; AAA46064.1; -.
DR EMBL: AJ004601; CA806087.1; -.
DR PIR: J01893; J01893.
KV Late protein: Phosphorylation; Trans-acting factor; Structural protein;
KW Late protein: Phosphorylation.
SQ SEQUENCE 739 AA; 80756 MW; 8F81232DF08F3DD CRC64;
Query Match 30.88; Score 46.5; DB 1; Length 739;
Best Local Similarity 63.28; Pred. No. 30;
Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 1;
OY 3 LRNDVALCLSPPLAARR 21
DB 669 LRLRPA---SPPLAKRR 684
RESULT 6
UL47_HSVBP STANDARD; PRT; 742 AA.
AC P30021;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE 80.7 kDa alpha trans-inducing protein (VP8 tegument protein) (VP7
DE protein) (107 kDa protein).
GN VP8.
OS Bovine herpesvirus type 1 (strain P8-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10324;
RN [1]

```

```

RP  SEQUENCE FROM N.A. PubMed=1662698;
RX  MEDLINE=9211550; Mista V.;
RT  "The most abundant protein in bovine herpes 1 virions is a homologue
RT  of herpes simplex virus type 1 UL47."
RL  J. Gen. Virol. 72:3077-3084(1991).
CC  -1- FUNCTION: MODULATOR OF ALPHA-TIF (VM65 PHOSPHOPROTEIN) TRANS-
CC  ACTIVATION. UL47 MAY HAVE KINASE ACTIVITY.
CC  -1- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
CC  -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
CC  -1- PTM: PHOSPHORYLATED.
CC  -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
CC  EBV-1 13, AND VZV 11.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D10327; BAA01170.1; -
DR  EMBL: Z11610; CA77683.1; -
DR  PIR: J01435; TNBEB1.
KW  Transcription regulation; Trans-acting factor; Structural protein;
KW  Late protein; Phosphorylation.
SQ  SEQUENCE 742 AA; 80744 MW; 85979D8C2C93C89 CRC64;

Query Match 30.8%; Score 46.5; DB 1; Length 742;
Best Local Similarity 63.2%; Pred. No. 30;
Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

OY  3 LRLNDVALCLSPPLARRR 21
DB  672 LRLRPVA--SPPLAGKR 687

RESULT 7
DSBB_NEIMA STANDARD; PRT; 162 AA.
ID  O9JTA9;
AC  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DR  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Disulfide bond formation protein B (Disulfide oxidoreductase).
GN  DSBB OR NMA1903.
OS  Neisseria meningitidis (serogroup A).
OC  Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX  NCBI_TaxID=65699;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX  MEDLIN=2022356; PubMed=10761919;
RA  Parthill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA  Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA  Davies K., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,
RA  Jags R., Leather S., Moule S., Mungall K., Quail M.A.,
RA  Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA  Whitehead S., Spratt B.G., Barrell B.G.;
RT  "Complete DNA sequence of a serogroup A strain of Neisseria
RT  meningitidis 22491."
RL  Nature 404:502-506(2000).
CC  -1- FUNCTION: REQUIRED FOR DISULFIDE BOND FORMATION IN SOME
CC  PERIPLASMIC PROTEINS. ACTS BY OXIDIZING THE DSBA PROTEIN (BY
CC  SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC  (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE DSBB FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its

```

```

CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AL162757; CAB85124.1; ALT_INT.
DR  InterPro: IPR003752; DSBB.
DR  Pfam: PF02600; DsbB; 1.
KW  Oxidoreductase; Redox-active center; Electron transport; Chaperone;
KW  Transmembrane; Inner membrane; Complete proteome.
FT  DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
FT  TRANSSEM 9 29 POTENTIAL.
FT  DOMAIN 30 35 PERIPLASMIC (POTENTIAL).
FT  TRANSSEM 36 56 POTENTIAL.
FT  DOMAIN 57 64 CYTOPLASMIC (POTENTIAL).
FT  TRANSSEM 65 85 POTENTIAL.
FT  DOMAIN 86 133 PERIPLASMIC (POTENTIAL).
FT  TRANSSEM 134 154 CYTOPLASMIC (POTENTIAL).
FT  DOMAIN 155 162 CYTOPLASMIC (POTENTIAL).
FT  DISULFID 35 38 REDOX-ACTIVE (BY SIMILARITY).
FT  DISULFID 101 128 REDOX-ACTIVE (BY SIMILARITY).
SQ  SEQUENCE 162 AA; 17601 MW; FFC0D64029A08802 CRC64;

Query Match 30.8%; Score 46; DB 1; Length 162;
Best Local Similarity 47.8%; Pred. No. 7.3;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY  9 ALCLSPPLARRRRSGRFVAV 31
DB  49 ALCAATVIMCPRKAGLFGAV 71

RESULT 8
PSBO_ONOVI STANDARD; PRT; 231 AA.
ID  O22591;
AC  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DR  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Oxygen-evolving enhancer protein 3, chloroplast precursor (OE3) (16
DE  kDa subunit of oxygen evolving system of photosystem II) (OE3 16 kDa
DE  subunit).
GN  PSBO.
OS  Onobrychis viciifolia (Common sainfoin).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids I; Fabales; Fabaceae; Papilionoideae; Hedyosaraceae;
OC  Onobrychis.
OX  NCBI_TaxID=3882;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Joseph R.G.;
RT  "cDNA encoding oxygen-evolving enhancer protein 3 precursor from
RT  Onobrychis viciifolia."
RL  Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC  -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC  with the photosystem II complex.
CC  -1- SIMILARITY: BELONGS TO THE PSBO FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AF026400; AAB81994.1; -
DR  Mendel: 25257; OnoVI:PSBQ;25257.
KW  Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW  Thylakoid; Membrane.
FT  TRANSIT 1 82 CHLOROPLAST (BY SIMILARITY).
FT  CHAIN 83 231 OXYGEN-EVOLVING ENHANCER PROTEIN 3.

```


Query Match 29.8%; Score 45; DB 1; Length 268;
 Best Local Similarity 34.5%; Pred. No. 17;
 Matches 10; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 3 LRNDVALCLSPPLARRRRSSGRVAV 31
 DB 164 IQVDTLAKTIQDPLAKLRERKSDRGVY 192

RESULT 12

YD3D_SCHPO STANDARD; PRT; 533 AA.

AC Q10277;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical RNA-binding protein C1367.13C in chromosome I.
 GN SPAC1367.13C OR SPAC663.01C.
 OS Schizosaccharomyces pombe (Fission yeast)
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 [1]
 RN SEQUENCE OF 1-394 FROM N.A.
 RC STRAIN=972;
 RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.,
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN SEQUENCE OF 363-533 FROM N.A.
 RC STRAIN=972;
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.,
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC -1- SIMILARITY: TO YEAST YH024W.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; Z69731; CA93614.1; -
 DR EMBL; Z69729; CA93601.1; -
 DR InterPro; IPR000504; RRM.
 DR Pfam; PF00076; RRM; 2.
 DR SMART; SMO0360; RRM; 2.
 DR PROSITE; PSS0102; RRM; 2.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 KW Hypothetical protein; RNA-binding; Repeat.
 FT DOMAIN 79
 FT 365 441 RNA-BINDING (RRM) 2.
 FT DOMAIN
 SQ SEQUENCE 533 AA; 59118 MW; A2ABA1633A0EC572 CRC64;

Query Match 29.8%; Score 45; DB 1; Length 533;
 Best Local Similarity 60.0%; Pred. No. 35;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 13 SPPLAARRRRSSGR 27
 DB 246 SPNNAHSRRRSQCK 260

RESULT 13

HYPF_RHOCA STANDARD; PRT; 739 AA.

AC Q02987;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hydrogenase maturation protein hypf.

GN HYPF.
 OS Rhodobacter capsulatus (Rhodospirillum rubrum capsulata).
 OS Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 CC Rhodobacter.
 OX NCBI_TaxID=1061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33303 / B10;
 RA MEDLINE=93268090; PubMed=8497190;
 RA Colbeau A., Richard P., Toussaint B., Caballero F.J., Elster C.,
 RA Delphin C., Smith R.L., Chabert J., Vignais P.M.,
 RT Organization of the genes necessary for hydrogenase expression in
 RT Rhodobacter capsulatus. Sequence analysis and identification of two
 RT hyp regulatory mutants. J. Biol. Chem. 268:11529-11533 (1993).
 RL Mol. Microbiol. 8:15-29(1993).
 RN [2]
 RP FUNCTION.

RA MEDLINE=98151233; PubMed=9492269;
 RA Colbeau A., Elsen S., Tomiyama M., Zorin N.A., Dimon B., Vignais P.M.,
 RT "Rhodobacter capsulatus HYPF is involved in regulation of hydrogenase
 RT synthesis through the HupV proteins." J. Biol. Chem. 273:7111-7116 (1998).
 RL Eur. J. Biochem. 251:65-71(1998).
 CC -1- FUNCTION: INVOLVED IN THE REGULATION OF HYDROGENASE EXPRESSION.
 CC -1- SIMILARITY: BELONGS TO THE HYPF FAMILY.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC ACYLPHOSPHATASE FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; Z15087; CA78795.1; -
 DR PIR; S25692; S25692.
 DR PIR; S32948; S32948.
 DR InterPro; IPR001792; Acylphosphatase.
 DR InterPro; IPR000666; Sna5_yjc10_yrhc.
 DR Pfam; PF00708; Acylphosphatase; 1.
 DR Pfam; PF01300; Sna5_yjc10_yrhc; 1.
 DR PRODOM; PD001884; Acylphosphatase; 1.
 DR PROSITE; PS00150; ACYLPHOSPHATASE_1; FALSE_NEG.
 KW Zinc-finger.
 FT DOMAIN 1
 FT 104 128 ACYLPHOSPHATASE-LIKE.
 FT 2N-FING 153 178 C4-TYPE (POTENTIAL).
 FT MTAGCN 723 739 MISSING: IN RS13 MUTANT; RESULTS IN
 FT LOSS OF HYDROGENASE ACTIVITY.
 SQ SEQUENCE 739 AA; 77878 MW; 19F3AD0C65D90B9 CRC64;

Query Match 29.8%; Score 45; DB 1; Length 739;
 Best Local Similarity 40.9%; Pred. No. 50;
 Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAURLNDVALCLSPPLAARRR 22
 DB 351 IARLIDSVYRVDPKYLRRAR 372

RESULT 14

POLG_HCV1 STANDARD; PRT; 3011 AA.

AC P26664;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
 DE Envelope glycoprotein E1 (GP32); Envelope glycoprotein E2
 DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
 DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepatitis C virus
 DE (EC 3.4.21.96); Nonstructural protein NS4A (P4); Nonstructural protein

DE NS4B (P27): Nonstructural protein NS5A (P56): Nonstructural protein
 DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
 OS Hepatitis C virus (isolate 1) (HCV).
 CC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 CC Hepacivirus.
 OX NCBI_TaxID=1104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91172826; PubMed=1848704;
 RA Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
 RA Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,
 RA Bradley D.W., Kuo G., Houghton M.;
 RT "Genetic organization and diversity of the hepatitis C virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPID-PROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M62321. AAA45676.1. -.
 DR PIR: A39166; GNMVC3.
 DR HSSP: P27958; IHEI.
 DR MEROPS: S29.001. -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR002518; HCV_NS2.
 DR InterPro: IPR004109; HCV_NS3.
 DR InterPro: IPR000745; HCV_NS4A.
 DR InterPro: IPR001490; HCV_NS4B.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRp.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF01543; HCV_capsid.1.
 DR Pfam: PF01542; HCV_core.1.
 DR Pfam: PF01539; HCV_env.1.
 DR Pfam: PF01560; HCV_NS1.1.
 DR Pfam: PF01538; HCV_NS2.1.
 DR Pfam: PF02907; HCV_NS3.1.
 DR Pfam: PF01006; HCV_NS4a.1.
 DR Pfam: PF01001; HCV_NS4b.1.
 DR Pfam: PF01506; HCV_NS5a.1.
 DR Pfam: PF00998; HCV_RdRp.1.
 DR Pfam: PF00271; Helicase_C.1.
 DR Pfam: PF016062; HCV_NS1.1.
 DR Pfam: PF016062; HCV_NS1.1.
 DR SMART: SM00492; Helicase.1.
 KM Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
 KM Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
 KM Transmembrane; Nonstructural
 FT INT_MBT 1
 FT CHAIN 1
 FT CHAIN 115
 FT CHAIN 116
 FT CHAIN 191
 FT CHAIN 192
 FT CHAIN 384
 FT CHAIN 729
 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
 NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).

FT CHAIN 730 1006 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
 FT CHAIN 1007 1615 PROTEASE/HELICASE NS3 (POTENTIAL).
 FT CHAIN 1616 1862 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
 FT CHAIN 1863 2013 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
 FT CHAIN 2014 3011 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
 FT TRANSMEM 347 369 POTENTIAL.
 FT ACT_SITE 1083 1083 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 1107 1107 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 1165 1165 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT NP_BIND 1230 1237 ATP (POTENTIAL).
 FT SITE 1316 1319 DECH BOX.
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 476 476 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 645 645 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2041 2041 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2077 2077 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2240 2240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2364 2364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2789 2789 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 3011 AA; 327197 MW; 65F8C9447FC5BA99 CRC64;
 Query Match 29.8%; Score 45; DB 1; Length 3011;
 Best Local Similarity 48.1%; Pred. No. 2.2e+02;
 Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;
 Qy 5 LNDVALCIS---PPLARRRRSSGR 27
 Db 2902 INRVACIRKGVPLAMHRRASVR 2928
 ID YEA1 YEA2 STANDARD; PRT; 246 AA.
 AC P53970;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Hypothetical 27.7 kDa protein in UME3-HDA1 intergenic region.
 GN YNL024C OR N2809.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Duesterhoft A., Floeth M., Filtz C., Heuss-Nelitzel D.,
 RA Hilbert H., Moestl D.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Andre B., Itagui Housaini I., Uristraza L.A., Viissers S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 11:08:45 ; Search time 27 Seconds
(without alignments)
198.624 Million cell updates/sec

Title: US-09-358-321c-15

Perfect score: 151

Sequence: 1 MALRLNDVALCLSPPLARRRRSSGRFVAV 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	37.1	358	2 Q93S04	Q93S04 pseudomonas
2	53	35.1	366	16 Q9KIL9	Q9KIL9 vibrio chol
3	51	33.8	679	16 Q914A6	Q914A6 pseudomonas
4	51	33.8	918	3 Q13663	Q13663 schizosacch
5	50	33.1	112	2 Q9KIC0	Q9KIC0 salmonella
6	50	33.1	208	16 Q9CMR1	Q9CMR1 pasteurella
7	50	33.1	242	16 Q84019	Q84019 chlamydia t
8	50	33.1	720	5 Q9YWA5	Q9YWA5 drosophila
9	49.5	32.6	221	11 Q64251	Q64251 mus musculu
10	49.5	32.6	221	11 Q9NM25	Q9NM25 mus musculu
11	49.5	32.6	1940	16 Q9RUC8	Q9RUC8 delnoccocus
12	49.5	32.5	258	2 Q9WX14	Q9WX14 streptomyce
13	49	32.5	498	10 Q9FVR3	Q9FVR3 citrus sine
14	49	32.5	609	10 Q9AYC4	Q9AYC4 oryza sativ
15	49	32.5	699	10 Q9MAH1	Q9MAH1 arabidopsis
16	48	31.8	1066	5 Q9TZK4	Q9TZK4 caenorhabdi

17	47	31.1	126	17 Q9A084	Q9A084 halobacteri
18	47	31.1	138	2 Q9A079	Q9A079 hordeum vul
19	47	31.1	228	10 Q9ANC3	Q9ANC3 bradyrhizob
20	47	31.1	276	10 Q23890	Q23890 oryza sativ
21	47	31.1	337	12 Q55826	Q55826 san perilla
22	47	31.1	404	10 Q94F85	Q94F85 oryza sativ
23	47	31.1	509	10 Q91J77	Q91J77 oryza sativ
24	47	31.1	672	10 Q9ZMB1	Q9ZMB1 arabidopsis
25	47	31.1	783	10 Q80558	Q80558 arabidopsis
26	47	31.1	832	5 Q95Q90	Q95Q90 caenorhabdi
27	47	31.1	835	5 Q95Q91	Q95Q91 caenorhabdi
28	47	31.1	862	10 Q9FTQ0	Q9FTQ0 oryza sativ
29	47	31.1	961	10 Q22937	Q22937 arabidopsis
30	47	31.1	1029	10 Q9SKA3	Q9SKA3 arabidopsis
31	47	31.1	1169	5 Q95Q07	Q95Q07 caenorhabdi
32	47	31.1	1819	5 Q16625	Q16625 caenorhabdi
33	47	31.1	1876	5 Q95Q08	Q95Q08 caenorhabdi
34	46.5	30.8	280	10 Q94JK8	Q94JK8 oryza sativ
35	46.5	30.8	450	5 Q9V806	Q9V806 drosophila
36	46.5	30.8	639	16 Q92Z13	Q92Z13 rhizobium m
37	46.5	30.8	2499	5 Q9N9N0	Q9N9N0 leishmania
38	46	30.5	779	10 Q9SD15	Q9SD15 oryza sativ
39	46	30.5	115	16 Q9K7C3	Q9K7C3 mycobacteri
40	46	30.5	122	12 Q98000	Q98000 hepatitis c
41	46	30.5	122	12 Q98002	Q98002 hepatitis c
42	46	30.5	122	12 Q97593	Q97593 hepatitis c
43	46	30.5	226	4 Q96BU4	Q96BU4 homo sapien
44	46	30.5	236	2 Q9A474	Q9A474 agrobacteri
45	46	30.5	297	2 Q9KIF0	Q9KIF0 streptomyce

ALIGNMENTS

RESULT 1									
Q93S04	PRELIMINARY:	PRT:	358 AA.						
AC Q93S04;									
DT 01-DEC-2001 (TREMBLrel. 19, Created)									
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)									
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)									
DE HOMOSERINE-ORTHO-ACYLTRANSFERASE.									
GN METX.									
OS Pseudomonas putida.									
OC Bacteria: Proteobacteria; gamma subdivision: Pseudomonadaceae;									
OC Pseudomonas.									
OX NCBI_TaxID=303;									
RN [1]									
RP SEQUENCE FROM N.A.									
RC STRAIN-DOT-TLE;									
RA Alaminos M., Ramos J.L.;									
RT "The methionine biosynthetic pathway for homoserine in P. putida									
RT involves the metW, metX, metZ, metH, and metE gene products.";									
RL Arch. Microbiol. 0:0-0(2001).									
DR EMBL: AY028975; AKA4978.1; -									
KW Transferase; Acyltransferase.									
SO SEQUENCE 358 AA; 38387 MW; 4D29C057EFPB837 CRC64;									
Query Match	37.1%;	Score 56;	DB 2;	Length 358;					
Best Local Similarity	48.1%;	Pred. No. 1.7;							
Matches 13;	Conservative	2;	Mismatches	6;	Indels	6;	Gaps	1;	
Qy 7	DVALCLSPPLA-----ARRRRSSGR	27							
Db 313	DVCICRSPPLALLSGPFARRRRACGR	339							
RESULT 2									
Q9KIL9	PRELIMINARY:	PRT:	366 AA.						
ID Q9KIL9									
AC Q9KIL9;									
DT 01-OCT-2000 (TREMBLrel. 15, Created)									
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)									

DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE MOLYBDENUM ABC TRANSPORTER, ATP-BINDING PROTEIN.
 GN VCA0724.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.,
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae."
 RL Nature 406:477-483(2000).
 DR EMBL: AE004401; AAF96623.1; -.
 DR TIGR: VCA0724; -.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 DR ATP-binding; Complete proteome.
 KW SEQUENCE 366 AA; 40664 MW; F1F3662AAFE7A65 CRC64;

Query Match 35.1%; Score 53; DB 16; Length 366;
 Best Local Similarity 38.7%; Pred. No. 5;
 Matches 12; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
 QY 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 31
 DB 276 MOVRANDYSIALDKPTASSIRNLIPARIVAI 306

RESULT 3
 Q914A6 PRELIMINARY; PRT; 679 AA.
 AC Q914A6;
 DT 01-MAR-2001 (TREMblrel. 15, Created)
 DT 01-MAR-2001 (TREMblrel. 15, last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, last annotation update)
 DE HYPOTHETICAL PROTEIN PA1232.
 GN PA1232.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OX NCBI_TaxID=287;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.C., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wagman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:958-964(2000).
 DR EMBL: AE004552; AAG04621.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 679 AA; 73496 MW; 0A3160805BA64671 CRC64;

Query Match 33.8%; Score 51; DB 16; Length 679;
 Best Local Similarity 43.3%; Pred. No. 18;
 Matches 13; Conservative 3; Mismatches 8; Indels 6; Gaps 1;

QY 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 30
 DB 620 IAVRHRD-----PPLAARRHARRAGARLAA 643
 RESULT 4
 O13663 PRELIMINARY; PRT; 918 AA.
 AC O13663;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-JUN-1998 (TREMblrel. 06, last annotation update)
 DE ORF YL031C.
 GN P1072 OR SPBC27B12.06
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-972 H-;
 RA Kuchida N., Yamazaki S., Tanaka T., Jinno K., Haikawa Y., Yamazaki J.,
 RA Yamamoto S., Sekine M., Oguchi A., Nagai Y., Sakai M., Aoki K.,
 RA Ogura K., Otsuka R., Kudo Y., Yanagida M., Machida M., Zhang M.Q.,
 RL submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN-972H-;
 RA Wood V., Rajendram M.A., Barrell B.G., Lauber J., Hilbert H.,
 RA Duesthoeft A.;
 RL submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB004539; BAA21454.1; -.
 DR EMBL: AL021766; CAA16901.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 918 AA; 104992 MW; EB6870E37BD0C314 CRC64;

Query Match 33.8%; Score 51; DB 3; Length 918;
 Best Local Similarity 45.5%; Pred. No. 25;
 Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
 QY 5 LNDVALCLSPPLAARRRRSSG 26
 DB 53 LNOTSLCENPPADYREKKNSSG 74

RESULT 5
 O9K1C0 PRELIMINARY; PRT; 112 AA.
 AC O9K1C0;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE HYPOTHETICAL 12.5 KDA PROTEIN.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OX NCBI_TaxID=602;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-140285;
 RX MEDLINE=20319062; PubMed=10861017;
 RA Miao E.A., Miller S.I.;
 RT "A conserved amino acid sequence directing intracellular type III
 RT secretion by Salmonella typhimurium."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:7539-7544(2000).
 DR EMBL: AF236075; AAF82083.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 112 AA; 12542 MW; 51018F2PE8573612 CRC64;

Query Match 33.1%; Score 50; DB 2; Length 112;
 Best Local Similarity 50.0%; Pred. No. 4.6;


```

RESULT 2
US-08-194-981E-17
; Sequence 17, Application US/08194981E
; Patent No. 5886157
; GENERAL INFORMATION:
; APPLICANT: GUENGERICH, F. Peter
; APPLICANT: GUO, Zuyu
; APPLICANT: SANDHU, Punam
; APPLICANT: GILLAM, Elizabeth M. J.
; TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF
; TITLE OF INVENTION: HUMAN
; TITLE OF INVENTION: CYTOCHROME P450
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P. C.
; STREET: Suite 1200, 127 Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,981E
; FILING DATE: February 10, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Selby
; REGISTRATION NUMBER: 38,298
; REFERENCE/DOCKET NUMBER: 22000.0022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; FRAGMENT TYPE: N-terminal
;
US-08-194-981E-17
;
Query Match          29.8%; Score 45; DB 2; Length 41;
Best Local Similarity 51.9%; Pred. No. 3.8;
Matches 14; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

```

```

CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREBSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-9556
; TELEFAX: 708-937-6365
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-867-611-14
;
Query Match          29.8%; Score 45; DB 4; Length 393;
Best Local Similarity 48.1%; Pred. No. 47;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

```

STATE: ILLINOIS
COUNTRY: U.S.
ZIP: 60065-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06965A
FILING DATE: 19920821
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834PC.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-06965A-19

Query Match 29.8%; Score 45; DB 5; Length 393;
Best Local Similarity 48.1%; Pred. No. 47;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 5 LNDVALCLSLPPLAARRRRSSGR 27
DB 284 INRVAACTLRKLGVPPLRAWRHRARSVR 310

RESULT 5
US-08-819-013-1
Sequence 1, Application US/08819013
Patent No. 5994522
GENERAL INFORMATION:
APPLICANT: Chan, Andrew C.
TITLE OF INVENTION: BINK PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSER: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,013
FILING DATE: 17-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/788,322
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 36,304
REFERENCE/DOCKET NUMBER: A-64383-1/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-8249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-819-013-1

Query Match 29.8%; Score 45; DB 2; Length 456;
Best Local Similarity 38.7%; Pred. No. 55;
Matches 12; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 31
DB 261 VASQONASVCEERPIPAERHRGSSHROEAV 291

RESULT 6
US-09-208-140-4
Sequence 4, Application US/09208140
Patent No. 6228576
GENERAL INFORMATION:
APPLICANT: Del Vecchio, Alfred
TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN
TITLE OF INVENTION: AND METHODS THEREOF TO IDENTIFY ANTIVIRAL COMPOUNDS
FILE REFERENCE: P50743
CURRENT APPLICATION NUMBER: US/09/208,140
CURRENT FILING DATE: 1998-12-09
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 590
TYPE: PRT
ORGANISM: Viral
US-09-208-140-4

Query Match 29.8%; Score 45; DB 4; Length 590;
Best Local Similarity 48.1%; Pred. No. 74;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 5 LNDVALCLSLPPLAARRRRSSGR 27
DB 502 INRVAACTLRKLGVPPLRAWRHRARSVR 528

RESULT 7
US-09-208-140-2
Sequence 2, Application US/09208140
Patent No. 6228576
GENERAL INFORMATION:
APPLICANT: Del Vecchio, Alfred
TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN
TITLE OF INVENTION: AND METHODS THEREOF TO IDENTIFY ANTIVIRAL COMPOUNDS
FILE REFERENCE: P50743
CURRENT APPLICATION NUMBER: US/09/208,140
CURRENT FILING DATE: 1998-12-09
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 591
TYPE: PRT
ORGANISM: Viral
US-09-208-140-2

Query Match 29.8%; Score 45; DB 4; Length 591;
Best Local Similarity 48.1%; Pred. No. 74;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 5 LNDVALCLSLPPLAARRRRSSGR 27
DB 482 INRVAACTLRKLGVPPLRAWRHRARSVR 508

RESULT 8

US-08-722-806A-2
; Sequence 2, Application US/08722806A
; Patent No. 5981247
; GENERAL INFORMATION:
; APPLICANT: Hagedorn, Curt H.
; APPLICANT: Al, Reinoldus H.
; * TITLE OF INVENTION: Recombinant Hepatitis C Virus RNA
; * TITLE OF INVENTION: Replicase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,806A
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004383
; FILING DATE: 27-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 76-95US
; TELECOMMUNICATION INFORMATION:
; * TELEPHONE: (303) 499-8080
; * TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-722-806A-2

Query Match 29.8%; Score 45; DB 2; Length 593;
Best Local Similarity 48.1%; Pred. No. 74;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 5 LNDVALCLIS---PPLAARRRRSSGR 27
DB 484 INRVACLRKLGVPPLRAMRHARSVR 510

RESULT 9
US-09-337-028-2
; Sequence 2, Application US/09337028
; Patent No. 6248589
; GENERAL INFORMATION:
; APPLICANT: Hagedorn, Curt H.
; APPLICANT: Al, Reinoldus H.
; * TITLE OF INVENTION: Recombinant Hepatitis C Virus RNA Replicase
; * TITLE OF INVENTION: Replicase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,806A
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004383
; FILING DATE: 27-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 76-95US
; TELECOMMUNICATION INFORMATION:
; * TELEPHONE: (303) 499-8080
; * TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-722-806A-2

US-09-337-028-2
Query Match 29.8%; Score 45; DB 4; Length 593;
Best Local Similarity 48.1%; Pred. No. 74;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 5 LNDVALCLIS---PPLAARRRRSSGR 27
DB 484 INRVACLRKLGVPPLRAMRHARSVR 510

RESULT 10
US-08-444-818-138
; Sequence 138, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; * TITLE OF INVENTION: NANBY Diagnostics and Vaccines
; * TITLE OF INVENTION: Replicase
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 359-3876
; TELEFAX: (508) 359-3885
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2995 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-444-818-138

Query Match 29.8%; Score 45; DB 4; Length 2995;
Best Local Similarity 48.1%; Pred. No. 4,6e+02;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

QY 5 LNDVALCLIS---PPLAARRRRSSGR 27
DB 2902 INRVACLRKLGVPPLRAMRHARSVR 2928

RESULT 11
US-08-453-552-2
; Sequence 2, Application US/08453552
; Patent No. 5667992
; GENERAL INFORMATION:
; APPLICANT: CASEY, JAMES M.
; APPLICANT: BODE, SUZANNE L.
; APPLICANT: ZECK, BILLY J.
; APPLICANT: YAMAGUCHI, JULIE
; APPLICANT: FRATL, DONALD E.

Query Match 29.8%; Score 45; DB 1; Length 3011;
Best Local Similarity 48.1%; Pred. No. 4.6e+02;
Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

Oy 5 LNDVALCLSL---PPLAARRRRSSGR 27
Db 2902 INRYAACLRKLGVPPLAARRRRASVR 2928

RESULT 14
US-07-910-760-10

; Sequence 10, Application US/07910760

; Patent No. 5683864

; GENERAL INFORMATION:

; APPLICANT: Houghton, Michael

; APPLICANT: Kuo, George

; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: P.O. Box 8097 (Int. Prop. R-440)

; CITY: Emeryville

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 94662-8097

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/910,760

; FILING DATE: 07-JUL-1992

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Blackburn Esq., Robert P.

; REGISTRATION NUMBER: 30,447

; REFERENCE/DOCKET NUMBER: 0101.002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (510) 601-2702

; TELEFAX: (510) 655-3542

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3011 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-910-760-10

Query Match 29.8%; Score 45; DB 1; Length 3011;

Best Local Similarity 48.1%; Pred. No. 4.6e+02;

Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

Oy 5 LNDVALCLSL---PPLAARRRRSSGR 27
Db 2902 INRYAACLRKLGVPPLAARRRRASVR 2928

RESULT 15
US-08-440-519-10

; Sequence 10, Application US/08440519

; Patent No. 5712087

; GENERAL INFORMATION:

; APPLICANT: Houghton, Michael

; APPLICANT: Kuo, George

; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: P.O. Box 8097 (Int. Prop. R-440)

CITY: Emeryville

STATE: CA

COUNTRY: U.S.A.

ZIP: 94662-8097

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/440,519

FILING DATE: 12-MAY-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/910,760

FILING DATE: 07-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Blackburn Esq., Robert P.

REGISTRATION NUMBER: 30,447

REFERENCE/DOCKET NUMBER: 0101.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 601-2702

TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 3011 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

Oy 5 LNDVALCLSL---PPLAARRRRSSGR 27

Db 2902 INRYAACLRKLGVPPLAARRRRASVR 2928

Query Match 29.8%; Score 45; DB 1; Length 3011;

Best Local Similarity 48.1%; Pred. No. 4.6e+02;

Matches 13; Conservative 1; Mismatches 9; Indels 4; Gaps 1;

Search completed: October 17, 2002, 11:14:32
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 10:28:40 : Search time 32 seconds

(without alignments)
107.603 Million cell updates/sec

Title: 05-09-358-321c-15

Perfect score: 151

Sequence: 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 31

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A: Geneseq.032802.*

1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*

2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*

3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*

4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*

5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*

6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*

7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*

8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*

9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*

10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*

11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*

12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*

13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*

14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*

15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*

16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*

17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*

18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*

19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*

20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*

21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*

22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	100.0	31	21	AAV68991
2	151	100.0	393	18	AAV09508
3	151	100.0	393	21	AAV68987
4	115	76.2	392	21	AAV32384
5	55	36.4	190	22	AAU42644
6	54	35.8	43	22	ABR40658
7	54	35.8	43	22	ABR24910
8	54	35.8	43	22	AAV61518
9	54	35.8	43	22	AAV74307
10	54	35.8	43	22	AAV34419
11	52	34.4	116	14	AAV35591

12	52	34.4	502	18	AAV45453	Arabidopsis thalia
13	50	33.1	720	22	ABR68640	Drosophila melanog
14	48.5	32.1	95	22	AAU45425	Propionibacterium
15	48	31.8	40	22	AAU12076	Human polypeptide
16	48	31.8	83	22	AAU61951	Propionibacterium
17	48	31.8	726	22	ABG10264	Novel human diago
18	48	31.8	992	20	AAW96265	Murine axin. Mus
19	47	31.1	55	22	AAU45543	Propionibacterium
20	47	31.1	129	21	AAV32776	Eucalyptus grandis
21	47	31.1	809	22	AAV41064	Human polypeptide
22	46.5	30.8	67	22	AAU40936	Propionibacterium
23	46.5	30.8	214	22	ABV10355	Human cDNA SEQ ID
24	46.5	30.8	214	22	AAU18001	Human immunoglobul
25	46.5	30.8	450	22	ABR62490	Drosophila melanog
26	46	30.5	38	22	ABR28830	Peptide #1481 enco
27	46	30.5	38	22	ABR34015	Peptide #1521 enco
28	46	30.5	38	22	ABR19456	Protein #1455 enco
29	46	30.5	38	22	AAV54783	Human brain expres
30	46	30.5	38	22	AAV67170	Human bone marrow
31	46	30.5	38	22	AAV15029	Peptide #1504 enco
32	46	30.5	38	22	AAV27467	Peptide #1504 enco
33	46	30.5	38	22	AAV02759	Peptide #1441 enco
34	46	30.5	203	22	ABV11632	Human dopamine rec
35	46	30.5	260	22	ABG01993	Novel human diago
36	46	30.5	389	21	AAV48414	Arabidopsis thalia
37	46	30.5	501	22	ABV71497	Drosophila melanog
38	46	30.5	3096	20	AAV06919	Hexaploid wheat PB
39	45	29.8	41	20	AAV93226	Human cytochrome P
40	45	29.8	70	22	AAV50290	Propionibacterium
41	45	29.8	87	21	AAV18539	Protein encoded by
42	45	29.8	112	22	AAU31466	Novel human secret
43	45	29.8	117	14	AAV35592	Tryptophan aporepr
44	45	29.8	118	22	ABV70823	Drosophila melanog
45	45	29.8	127	22	ABV71799	Drosophila melanog

ALIGNMENTS

RESULT 1

AAV68991 standard; Peptide: 31 AA.

XX AAV68991:

XX 30-MAY-2000 (first entry)

XX

XX Amino acid sequence of a maize delta9-desaturase fragment.

XX

XX Delta9-desaturase; antibody; transit peptide; passenger protein;

XX plant cell organelle; maize; stearyl-ACP-delta9-desaturase;

XX transgenic plant.

XX

XX Zea mays.

XX

XX WO200005391-A1.

XX

XX 03-FEB-2000.

XX

XX 21-JUL-1999; 99WO-US16405.

XX

XX 21-JUL-1998; 98US-0093587.

XX

XX (DOWC) DOW AGROSCIENCES LLC.

XX

XX Sukhplinda K, Hasler JM, Petell JK, Strickland JA, Folkerts O;

XX WPI; 2000-182711/16.

XX

XX Novel nucleic acid construct for down-regulating steady state levels of

XX proteins in plant cells, transgenic plants and their progeny

XX

XX Claim 22; Page 87; 114pp; English.

XX The present sequence represents a Zea mays (maize) delta9-desaturase
 CC fragment. The specification describes a construct encoding an antibody
 CC that can bind a transit peptide that directs an associated passenger
 CC protein to a plant cell organelle. The transit peptide sequence of
 CC the maize stearoyl-ACP-desaturase (delta9-desaturase) was
 CC determined, and used to produce antibodies of the invention. These
 CC antibodies were produced in transgenic plants of the invention. The
 CC constructs of the invention are useful for producing antibodies
 CC which decrease steady state levels of passenger proteins in the
 CC organelles of plant cells and plants, by binding to the transit
 CC peptide. This results in the production of transgenic plants which
 CC have altered steady state passenger protein levels.

XX
 CC Sequence 31 AA:

Query Match 100.0%; Score 151; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.6e-15;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MALRLNDVALCLSPPLAARRRRSSGREFVAV 31
 Db 1 MALRLNDVALCLSPPLAARRRRSSGREFVAV 31

RESULT 2
 AA09508
 ID AA09508 standard; Protein; 393 AA.
 AC AA09508;
 XX
 XX 16-JUL-1999 (first entry)
 DE Zea mays delta-9 desaturase.
 XX
 XX Maize; corn; Zea mays; delta-9 desaturase; GSSS; target; substrate;
 KW granule bound starch synthase; hammerhead ribozyme; hairpin ribozyme;
 KW modulation; gene expression; transgenic plant; cleavage; canola plant;
 KW caffeine synthesis; coffee plant; nicotine production; tobacco;
 KW fruit ripening; flower pigmentation; lignin production.
 XX
 XX Zea mays.
 OS
 XX WO9710328-A2.
 PN
 XX 20-MAR-1997.
 PD
 XX 12-JUL-1996; 96WO-US11669.
 PF
 XX 13-JUL-1995; 95US-0001135.
 PR
 XX (DOMC) RIBOZYME PHARM INC.
 PA (RIBO-) RIBOZYME PHARM INC.
 XX
 XX Edington BE, Folkeris O, Guo L, McSwigen JA, Merlo DJ;
 PI Merlo PAO, Skokut TA, Young SA, Zwick MG;
 DR WPI: 1997-202224/18.
 DR N-PSDB; AA62127.
 XX
 XX Ribozyme which modulates plant gene expression - preferably
 PT modulates expression of DELTA-9 desaturase or granule bound starch
 PT synthase in maize or canola
 PS
 PS Claim 37; Fig 1; 155pp; English.

CC The present invention describes an enzymatic nucleic acid molecule (I)
 CC with RNA cleaving activity, which modulates the expression of a plant
 CC gene. Also described is a gene comprising a cDNA sequence encoding maize
 CC Delta-9 desaturase. (I) can be used to modulate expression of a gene,
 CC preferably Delta-9 desaturase or a granule bound starch synthase (GSSS)
 CC gene, in a plant (preferably a maize or canola plant). (I) can be used
 CC to modulate caffeine synthesis in a coffee plant, nicotine production in

CC a tobacco plant, fruit ripening processes in an apple, tomato, pear,
 CC plum or peach plant, flower pigmentation in a rose, petunia,
 CC chrysanthemum or marigold plant or lignin production in a tobacco,
 CC aspen, poplar or pine plant. The present sequence represents the maize
 CC Delta-9 desaturase.

XX
 CC Sequence 393 AA:

Query Match 100.0%; Score 151; DB 18; Length 393;
 Best Local Similarity 100.0%; Pred. No. 3.9e-14;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MALRLNDVALCLSPPLAARRRRSSGREFVAV 31
 Db 1 MALRLNDVALCLSPPLAARRRRSSGREFVAV 31

RESULT 3
 AA68987
 ID AA68987 standard; Protein; 393 AA.
 AC AA68987;
 XX
 XX 30-MAY-2000 (first entry)
 DE Amino acid sequence of a maize delta9-desaturase.
 XX
 XX Delta9-desaturase; antibody; transit peptide; passenger protein;
 KW plant cell organelle; maize; stearoyl-ACP-desaturase;
 KW transgenic plant.
 XX
 XX Zea mays.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..31
 FT /note="signal peptide"
 FT Protein 32..393
 FT /note="mature protein"
 XX
 XX WO200005391-A1.
 PN
 XX 03-FEB-2000.
 PD
 XX 21-JUL-1999; 99WO-US16405.
 PF
 XX 21-JUL-1998; 98US-0093587.
 PR
 XX (DOMC) DOW AGROSCIENCES LLC.
 PA
 XX Sukhapiinda K, Hasler JM, Petell JK, Strickland JA, Folkeris O;
 PI WPI: 2000-182711/16.
 DR N-PSDB; AA61026.
 DR
 XX Novel nucleic acid construct for down-regulating steady state levels of
 PT proteins in plant cells, transgenic plants and their progeny
 PT
 PS
 PS Claim 22; Page 81-83; 114pp; English.

CC The present sequence represents a Zea mays (maize) delta9-desaturase.
 CC The specification describes a construct encoding an antibody that
 CC can bind a transit peptide that directs an associated passenger
 CC protein to a plant cell organelle. The transit peptide sequence of
 CC the maize stearoyl-ACP-desaturase (delta9-desaturase) was
 CC determined, and used to produce antibodies of the invention. These
 CC antibodies were produced in transgenic plants of the invention. The
 CC constructs of the invention are useful for producing antibodies
 CC which decrease steady state levels of passenger proteins in the
 CC organelles of plant cells and plants, by binding to the transit
 CC peptide. This results in the production of transgenic plants which
 CC have altered steady state passenger protein levels.

XX
 CC Sequence 393 AA;

```

Query Match          100.0%; Score 151; DB 21; Length 393;
Best Local Similarity 100.0%; Pred. No. 3,9e-14;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 31
   |||||
DB 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 31

RESULT 4
AAY32384
ID AAY32384 standard; Protein; 392 AA.
XX
AC AAY32384;
XX
DT 13-MAR-2000 (first entry)
XX
DE Corn delta-9 stearyl-ACP desaturase.
XX
KW Corn; maize; transgenic plant; lipid; food; feedstuff;
KW vegetable oil; seed oil; stearic acid; fatty acid desaturase;
KW delta-9 stearyl-ACP desaturase.
XX
OS Zea mays.
XX
PN M09964579-A2.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99MO-US12864.
XX
PR 11-JUN-1998; 98US-0088987.
XX
PA (DPO) DU PONT DE NEMOURS & CO E. I.
XX
PI Shen JB;
XX
DR WPI: 2000-097535/08.
DR N-PSDB; AAZ35179.
XX
PT New maize oleosin promoter, used for producing transgenic plants with
PT altered fatty acid composition of the oil, used e.g. in animal feeds -
XX
PS Example 3; Page 87-89; 108pp; English.
XX
CC This sequence represents the amino acid sequence of corn delta-9
CC stearyl-ACP desaturase, as deduced from an isolated cDNA clone (see
CC AAZ35179). The invention generally relates to the preparation and use
CC of nucleic acid fragments comprising all, or substantially all, of
CC a corn oleosin promoter (see AAZ35165-77), a stearyl-ACP desaturase
CC and a delta-12-desaturase (see AAZ35178), which can be used
CC individually or in combination to modify the lipid profile of corn.
CC A chimeric gene comprising the present nucleic acid, or the reverse
CC complement of it, operably linked to suitable regulatory sequences
CC is claimed. Expression of the chimeric gene results in an altered
CC corn stearic acid phenotype. Also claimed are seeds of the transgenic
CC plants, oil obtained from the grain of such plants, animal feed,
CC use of the oil in food, feed, and cooking oil or industrial
CC applications.
XX
SQ Sequence 392 AA;
XX
Query Match          76.2%; Score 115; DB 21; Length 392;
Best Local Similarity 83.9%; Pred. No. 8e-09;
Matches 26; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 31
   |||||
DB 1 MALRLNDVALCLSPPLAARRRRSSGRFVAV 29

RESULT 5

```

```

AAU42644
ID AAU42644 standard; Protein; 190 AA.
XX
AC AAU42644;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #3540.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN W0200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001MO-US12865.
XX
PR 21-APR-2000; 2000US-198647P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonmeuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
DR N-PSDB; AAS59518.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 3839; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 190 AA;
XX
Query Match          36.4%; Score 55; DB 22; Length 190;
Best Local Similarity 41.5%; Pred. No. 2.7;
Matches 17; Conservative 2; Mismatches 8; Indels 14; Gaps 2;
QY 2 ALRLNDVALC-----LSPLAARRRRSSGRF 28
   |||||
DB 104 ALRLRRRACSAHSCPIGADHLRPPMAAOSRRRRRHPGR 144

RESULT 6

```

```

ABBA0658
ID ABB40658 standard; Peptide: 43 AA.
AC ABB40658;
XX
XX 04-FEB-2002 (first entry)
DT
XX
DE Peptide #8164 encoded by human foetal liver single exon probe.
XX
XX Human: foetal liver; gene expression: single exon nucleic acid probe.
XX
OS Homo sapiens.
PN
XX WO200157277-A2.
PD
XX 09-AUG-2001.
PF
XX 30-JAN-2001: 2001WO-US00669.
PR 04-FEB-2000: 2000US-0180312.
PR 26-MAY-2000: 2000US-0207456.
PR 30-JUN-2000: 2000US-0608408.
PR 03-AUG-2000: 2000US-063366.
PR 21-SEP-2000: 2000US-0234687.
PR 27-SEP-2000: 2000US-0236359.
PR 04-OCT-2000: 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI: 2001-483447/52.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 27; SEQ ID NO 33293; 639bp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 43 AA:
SQ
Query Match 35.8%; Score 54; DB 22; Length 43;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0
QY 1 MALRLNDVALCLSPPLAARRR 21
   | : ||| ||| |
Db 15 MSLSPEKAVCVSPPLAQRTTR 35
RESULT 7
ABB24910
ID ABB24910 standard; Protein: 43 AA.
XX
XX ABB24910;
AC
XX 23-JAN-2002 (first entry)
DT
XX
DE Protein #6909 encoded by probe for measuring heart cell gene expression.
XX
XX Human: gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
```

```

XX Homo sapiens.
XX
XX W0200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001, 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0808408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX
XX WPI: 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX
XX Claim 15; SEQ ID No 26680; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX AB521535; AB441105). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 43 AA:
XX
XX Query Match 35.8%; Score 54; DB 22; Length 43;
XX Best Local Similarity 52.4%; Pred. No. 0.77;
XX Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 MAIRLNDVALCSPPLAARR 21
XX ||| :||:|||||
XX
XX Db 15 MSLASPEKAYCVSPPLAORTR 35
XX
XX RESULT 8
XX AAM61518
XX ID AAM61518 standard; Protein; 43 AA.
XX
XX AAM61518;
XX
XX AC
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 33623.
XX
XX Human: brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX W0200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001, 2001WO-US00667.
XX

```

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-483446/52.
 DR
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 33623; 650bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 CC
 SQ Sequence 43 AA;
 XX
 Query Match 35.8%; Score 54; DB 22; Length 43;
 Best Local Similarity 52.4%; Pred. No. 0.77;
 Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MALRLNDVALCLSPPLAARR 21
 ||| : ||:||||| |
 Db 15 MSLASPEKAVCVSPPLAQRTR 35
 RESULT 9
 AAM74307
 ID AAM74307 standard; Protein; 43 AA.
 XX
 AC AAM74307;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34613.
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34613.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-488900/53.
 DR
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 34613; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 CC
 SQ Sequence 43 AA;
 XX
 Query Match 35.8%; Score 54; DB 22; Length 43;
 Best Local Similarity 52.4%; Pred. No. 0.77;
 Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MALRLNDVALCLSPPLAARR 21
 ||| : ||:||||| |
 Db 15 MSLASPEKAVCVSPPLAQRTR 35
 RESULT 10
 AAM34419
 ID AAM34419 standard; Protein; 43 AA.
 XX
 AC AAM34419;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #8456 encoded by probe for measuring placental gene expression.
 DE
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-48897/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 27; SEQ ID NO 34688; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP:
 CC see A113315-A157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 CC
 SQ Sequence 43 AA;
 XX
 Query Match 35.8%; Score 54; DB 22; Length 43;

PF 23-MAR-2001; 2001MO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL12743.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 32712; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 720 AA;
Query Match 33.1%; Score 50; DB 22; Length 720;
Best Local Similarity 43.5%; Pred. No. 60;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 9 ALCISPLAARRRRSSGRVAV 31
DB 350 SLCSPPMNSKRUSKYSRDYDAV 372
RESULT 14
AAU45425
ID AAU45425 standard; Protein; 95 AA.
XX
XX AAU45425;
AC
XX 27-FEB-2002 (first entry)
DT
XX
XX Propionibacterium acnes immunogenic protein #6321.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX
XX MO200181561-A2.
PN
XX
XX 01-NOV-2001.
PD
XX
XX 20-APR-2001; 2001MO-US12865.
PF
XX
XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX

DR WPI; 2001-616774/71.
DR N-PSDB; AAS59525.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX
PS Example 1; SEQ ID NO 6620; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system; however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 95 AA;
Query Match 32.1%; Score 48.5; DB 22; Length 95;
Best Local Similarity 38.6%; Pred. No. 12;
Matches 17; Conservative 3; Mismatches 7; Indels 17; Gaps 3;
QY 2 ALRINDVAL-----CL-----SPPLAARRRRSSG-RF 28
DB 30 ALRIRDYGMPPRRRGRCGLVMRRDSSNPPGNSRRRRVGRPF 73
RESULT 15
AAO12076
ID AAO12076 standard; Protein; 40 AA.
XX
XX AAO12076;
AC
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 25968.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX
XX MO200164835-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001MO-US04927.
PF
XX
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
PR
XX
XX (HXSE-) HXSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-514838/56.
DR N-PSDB; AAI92007.
DR

XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

PS Claim 20: SEQ ID NO 25968; 1399bp + Sequence Listing; English.

XX
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA033910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 40 AA:

Query Match 31.88; Score 48; DB 22; Length 40;
 Best Local Similarity 83.38; Pred. No. 5.5;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY . 12 LSPPLARRRR 23
 | | | | | | | |
 DB 28 LQPPPARRRRR 39

Search completed: October 17, 2002, 11:11:13
 Job time : 33 secs